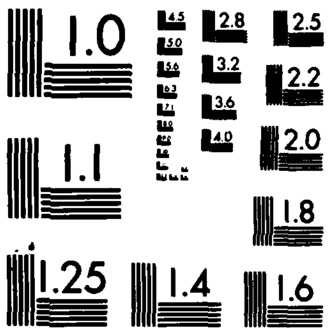
AD-A140 931 STATUS REPORT ON ASYMPTOTIC METHODS FOR THE ANALYSIS ESTIMATION AND CONTR..(U) MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR INFORMATION AND D.. 1/1

UNCLASSIFIED A S WILLSKY ET AL. JAN 84 LIDS-SR-1355 F/G 12/1 NL




MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

LIDS-SR- 1355

LABORATORY FOR INFORMATION AND DECISION SYSTEMS
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139



Status Report on

# ASYMPTOTIC METHODS FOR THE ANALYSIS, ESTIMATION, AND CONTROL OF STOCHASTIC DYNAMIC SYSTEMS

Grant AFOSR-82-0258

Covering the Period
January 1, 1983 to December 31, 1983

Prepared by
Professor Alan S. Willsky
Professor George C. Verghese



Submitted to: Dr. John Burns
Program Advisor
Directorate of Mathematical
and Information Sciences
Air Force Office of Scientific Research
Building 410
Bolling Air Force Base
Washington, D.C. 20332

January 1984

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

# REPORT DOCUMENTATION PAGE

| | |
|---|---|
| 1a. REPORT SECURITY CLASSIFICATION<br>UNCLASSIFIED | 1b. RESTRICTIVE MARKINGS |
| 2a. SECURITY CLASSIFICATION AUTHORITY | 3. DISTRIBUTION/AVAILABILITY OF REPORT<br>Approved for public release; distribution unlimited. |
| 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE | |
| 4. PERFORMING ORGANIZATION REPORT NUMBER(S) | 5. MONITORING ORGANIZATION REPORT NUMBER(S)<br>AFOSR-TR- 84-0401 |
| 6a. NAME OF PERFORMING ORGANIZATION<br>Massachusetts Institute of Technology / 6b. OFFICE SYMBOL (If applicable) | 7a. NAME OF MONITORING ORGANIZATION<br>Air Force Office of Scientific Research |
| 6c. ADDRESS (City, State and ZIP Code)<br>Laboratory for Information & Decision Systems<br>Cambridge MA 02139 | 7b. ADDRESS (City, State and ZIP Code)<br>Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332 |
| 8a. NAME OF FUNDING/SPONSORING ORGANIZATION<br>AFOSR / 8b. OFFICE SYMBOL (If applicable)<br>NM | 9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER<br>AFOSR-82-0258 |
| 8c. ADDRESS (City, State and ZIP Code)<br>Bolling AFB DC 20332 | 10. SOURCE OF FUNDING NOS.<br>PROGRAM ELEMENT NO. 61102F / PROJECT NO. 2304 / TASK NO. A1 / WORK UNIT NO. |

11. TITLE (Include Security Classification)
ASYMPTOTIC METHODS FOR THE ANALYSIS, ESTIMATION, AND CONTROL OF STOCHASTIC DYNAMIC SYSTEMS

12. PERSONAL AUTHOR(S)
Alan S. Willsky

| 13a. TYPE OF REPORT | 13b. TIME COVERED | 14. DATE OF REPORT (Yr., Mo., Day) | 15. PAGE COUNT |
|---|---|---|---|
| Interim | FROM 1/1/83 TO 31/12/83 | JAN 84 | 30 |

16. SUPPLEMENTARY NOTATION

| 17. COSATI CODES | | | 18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) |
|---|---|---|---|
| FIELD | GROUP | SUB. GR. | |
| | | | |


19. ABSTRACT (Continue on reverse if necessary and identify by block number)
In this report, the author presents a brief description of the research carried out by faculty, staff, and students in the Massachusetts Institute of Technology Laboratory for Information and Decision Systems under Grant AFOSR-82-0258. The principal investigator for this research is Professor Alan S. Willsky, and Professor George C. Verghese serves as a senior researcher on the project. The time period covered in this status report is from January 1, 1983 to December 31, 1983. The basic scope of this grant is to carry out fundamental research in the analysis, control, and estimation of complex systems, with particular emphasis on the use of methods of asymptotic analysis and multiple time scales to decompose complex problems into interconnections of simpler ones. During the time period covered by this report, significant progress has been made in several areas, leading to important results and to promising directions for further research. The specific topics covered in this report are: I. Analysis of Systems Possessing Multiple Time (CONTINUED)


| | |
|---|---|
| 20. DISTRIBUTION/AVAILABILITY OF ABSTRACT<br>UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐ | 21. ABSTRACT SECURITY CLASSIFICATION<br>UNCLASSIFIED |
| 22a. NAME OF RESPONSIBLE INDIVIDUAL<br>Dr. John A. Burns | 22b. TELEPHONE NUMBER (Include Area Code)<br>(202) 767-5028 / 22c. OFFICE SYMBOL<br>NM |

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE

ITEM #19, ABSTRACT, CONTINUED: Scales; and II. Control and Estimation. A list of publications supported by this grant is also included. The author refers heavily to these papers and reports for detailed technical developments.

## SUMMARY


In this report we present a brief description of the research carried out by faculty, staff, and students in the M.I.T. Laboratory for Information and Decision Systems under Grant AFOSR-82-0258. The principal investigator for this research is Professor Alan S. Willsky, and Prof. George C. Verghese serves as a senior researcher on the project. The time period covered in this status report is from January 1, 1983 to December 31, 1983.

The basic scope of this grant is to carry out fundamental research in the analysis, control, and estimation of complex systems, with particular emphasis on the use of methods of asymptotic analysis and multiple time scales to decompose complex problems into interconnections of simpler ones. During the time period covered by this report, significant progress has been made in several areas, leading to important results and to promising directions for further research.

The specific topics covered in this report are:

9 ~~I.~~ Analysis of Systems Possessing Multiple Time Scales. and

9 ~~II.~~ Control and Estimation .

A list of publications supported by this grant is also included. We refer heavily to these papers and reports for detailed technical developments.

## I. Analysis of Systems Possessing Multiple Time Scales

During the past year we have made significant progress in this portion of our research. Our recent work in this area is described in more detail in [7], [9], [11]. This work has focused on the examination of systems of the form

$$\dot{x}(t) = A(\varepsilon)x(t) \tag{1.1}$$

where $\varepsilon$ is a small parameter. In our earlier work [1], [4] we had developed a method for determining when it is possible to construct and for constructing an approximation to the solution of (1.1), of the form

$$x(t) = e^{A_0 t} e^{A_1 \varepsilon t} \cdots e^{A_r \varepsilon^r t} x(0) + O(\varepsilon) \tag{1.2}$$

where $A_i A_j = 0$ if $i \neq j$ and (1.2) is uniformly valid on $0 \leq t < \infty$. Thus (1.2) says that one can construct an $\varepsilon$-independent similarity transformation $\xi = Tx$ which puts the system (1.1) into a block-decoupled form which explicitly decomposes the system into subsystems evolving at different time-scales -- i.e.

$$\xi(t) = \text{diag}\,(e^{F_o t}, e^{F_1 \varepsilon t}, \cdots, e^{F_r \varepsilon^r t})\xi(0) + O(\varepsilon) \tag{1.3}$$

The motivation for this line of research came from a desire to extend the results and approaches developed by others, including Kokotovic, Campbell, O'Malley, Chow, Hoppensteadt, Khalil and Sannuti, who have demonstrated the utility of time-scale decompositions for system analysis, approximation, control, and estimation. The starting point for much of this work is a model of the form of (1.1) but in which the time-scale structure is explicitly displayed. For example, consider the two-time-scale model studied extensively by Kokotovic and co-workers:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ \varepsilon A_{21} & \varepsilon A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{1.4}$$

(The form actually used in many of these other works can be brought to (1.4) by a simple change of time-scale). For this model there is an extremely straightforward procedure for determining if a two-time-scale decomposition exists and, if so, for computing the fast and slow parts of the dynamics -- i.e. for constructing the similarity transformation which puts the system into the form of (1.3) (with $r = 2$). From this perspective we can view the contribution of [1] as determining the existence of and constructing such transformations when the time-scale behavior of the system is not evident by inspection. The price that was paid in [2], however, was a rather complex procedure involving nested projections and pseudo-inversions.

Our work [7], [9] in the past year has had as its goal the development of a methodology for analyzing (1.1) which combines the generality of [1] with the intuitive and algebraically simple results associated with models such as (1.4). We have now done this, and not only does this provide a very clear connection between our earlier work and the work of others, but it also establishes an algebraic framework for examining numerous other problems involving systems with multiple time scales.

The basis for our approach is to view $A(\varepsilon)$ as a matrix over the (local) ring $W$ of functions of $\varepsilon$ that are analytic at $\varepsilon = 0$. The matrix $A(\varepsilon)$ then has a Smith decomposition

$$A(\varepsilon) = P(\varepsilon)D(\varepsilon)Q(\varepsilon) \tag{1.5}$$

where $P(\varepsilon)$ and $Q(\varepsilon)$ are unimodular, i.e. $|P(0)| \neq 0$, $|Q(0)| \neq 0$ and

$$D(\varepsilon) = \text{diag}\,(\varepsilon^{i_1} I_1, \cdots, \varepsilon^{i_m} I_m) \tag{1.6}$$

where $0 \leq i_1 < \cdots < i_m$, and the identity matrices $I_j$ may have different dimensions. (We have assumed here that $A(\varepsilon)$ is nonsingular on some interval of the form $(0, \varepsilon_0)$; in the more general case $D(\varepsilon)$ would also include a zero diagonal block). The diagonal elements of $D(\varepsilon)$ are the invariant factors of $A(\varepsilon)$.

As described in [7], because of the unimodularity of $P(\varepsilon)$, one can show that a uniform approximation of $x(t)$ with the same time scale structure as (1.1) is $P(0)z(t)$, where

$$\dot{z} = D(\varepsilon)\bar{A}z \quad , \quad \bar{A} = Q(0)P(0) \tag{1.7}$$

Note that this system is in a form that is the natural generalization of (1.4):

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \vdots \\ \dot{z}_m \end{bmatrix} = \begin{bmatrix} \varepsilon^{i_1}A_{11} & \varepsilon^{i_1}A_{12} & \cdots & \varepsilon^{i_1}A_{1m} \\ \varepsilon^{i_2}A_{21} & \varepsilon^{i_2}A_{22} & \cdots & \varepsilon^{i_2}A_{2m} \\ \vdots & \vdots & & \\ \varepsilon^{i_m}A_{m1} & \varepsilon^{i_m}A_{m2} & \cdots & \varepsilon^{i_m}A_{mm} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} \tag{1.8}$$

Consequently there is a relatively straightforward procedure -- involving successive Schur complements of $\bar{A}$ -- for determining if (1.8) (and hence (1.1)) has well-behaved time scales, and for constructing the transformation which brings (1.8) (and thus (1.1)) into diagonalized time-scale form (as in (1.3)).

This result is, in our opinion, of great significance, as it makes clear the essential algebraic nature of the problem of time-scale decompositions. In particular, it establishes the fact that the invariant factors of $A(\varepsilon)$ determine the time scale structure of (1.1). This opens the door to the detailed examination of numerous other problems, some of which we have begun to examine and others which wait for the near future. Some of these are as follows:

(1) Important issues in the design of feedback control for complex systems are the way in which feedback couplings should be structured and the effect such couplings have on overall system performance. Such effects can be quite dramatic, as has been documented in numerous examples of control designs based on reduced-order models (neglecting, for example, fast dynamics) which lead to closed-loop instabilities or severe performance degradations. A natural question to ask in the context we have described is the effect of feedback on time-scale structure. From our results we see that a key question is that of invariant factor assignment -- i.e. the changing of invariant factors (and hence time scales) by feedback. We have obtained some important results in this area, but we defer discussion of these to the following section.

(2) The key computational aspect of our time scale decmoposition procedure is the determination of the Smith decomposition (1.5), (1.6). We have been able to relate this computation explicitly to the projections and pseudo- inversions in our earlier method [2]. Procedures exist for the computation of Smith forms,[†] and we have begun to examine the implementation of such algorithms for the numerical computation of time

---

[†] P. Van Dooren, P. Dewilde, and J. Vandewalle, "On the Determination of the Smith-McMillan Form of a Rational Matrix from Its Laurent Expansion," IEEE Trans. Circuits and Systems, March 1979.

G. Verghese and T. Kailath, "Rational Matrix Structure," IEEE Trans. Aut. Control, Vol. AC-26, No. 2, pp. 434-439, April 1981.

scale decompositions of systems of the form of (1.1).

(3) The form (1.5), (1.6) provides us with the basis for answering another question we had posed in our proposal for this project. Specifically, consider the problem of characterizing matrices $A_1(\varepsilon)$ and $A_2(\varepsilon)$ that have identical time-scale behavior -- i.e. they both have the same approximation in (1.2) -- so that the difference $A_2(\varepsilon) - A_1(\varepsilon)$ is a regular perturbation of $A_1(\varepsilon)$. Clearly one necessary condition is

$$D_1(\varepsilon) = D_2(\varepsilon) \tag{1.9}$$

This, together with the condition

$$P_1(0) = P_2(0), \quad Q_1(0) = Q_2(0)$$

form a set of sufficient conditions, but the latter is not necessary because of the nonuniqueness of P and Q in the Smith decomposition. More generally one could allow

$$P_1(0) = P_2(0)R \quad , \quad Q_1(0) = R^{-1}A_2(0) \tag{1.10}$$

where R is a block-diagonal similarity transformation with blocks of sizes equal to those in $D_1(\varepsilon)$, but even this is not a necessary condition, since one can essentially add slow modes to faster ones without affecting asymptotic behavior. For example,

$$A_1(\varepsilon) = \begin{bmatrix} -1 & 0 \\ 0 & -\varepsilon \end{bmatrix}$$

and

$$A_2(\varepsilon) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -\varepsilon \end{bmatrix} = \begin{bmatrix} -1 & -\varepsilon \\ 0 & -\varepsilon \end{bmatrix}$$

have the property that

$$\lim_{\varepsilon\downarrow 0}\ \sup_{t\geq 0}\ \left|\left|e^{A_1(\varepsilon)t} - e^{A_2(\varepsilon)t}\right|\right| = 0$$

This suggests that upper-block diagonal transformations (with identities along the diagonal) on the right of P do not affect asymptotic equivalence, and the same is true of lower-block diagonal transformations (with identities along the diagonal) on the left of Q. We are presently completing the verification of necessary and sufficient conditions along these lines (and also allowing transformations as in (1.10)) for our notion of asymptotic equivalence.

(4) The fact that our method yields an algebraic connection between (1.1) and the explicit form of (1.7) provides us with a basis for evaluating bounds on convergence rates of time-scale approximations. Specifically, we have begun to investigate the construction of such bounds in terms of $||P(\varepsilon) - P(0)||$, $||Q(\varepsilon) - Q(0)||$, and the Schur complement structure of $Q(0)P(0)$. The work described in (6) to follow on higher order corrections to the time-scale approximation (see (1.19) - (1.22)) will also be of value here in providing a method for pinpointing the lead-order error terms.

(5) Points (3) and (4) above are extremely important in shedding light on theoretical problems in practical time-scale decomposition. In particular, it is rarely the case that a system is given in the form (1.1) with a small parameter identified. Rather, one is interested in starting with a system in the form

$$\dot{x}(t) = Fx(t) \ , \tag{1.11}$$

identifying a small number $\varepsilon$ and writing

$$F = \sum_{i=0}^{\infty} B_i \varepsilon^i \qquad (1.12)$$

so that the resulting time-scale approximation based on this representation is accurate to within some prescribed tolerance. This problem raises numerous issues. For example, which small elements of F should be viewed as order 1, as order $\varepsilon$, as order $\varepsilon^2$, etc.? Intuitively one must have that $||P(\varepsilon) - P(0)||$ is sufficiently smaller than the minimum singular value of $P(0)$ (similarly for $Q(\varepsilon)$ and $Q(0)$) so that it is fair to view the difference $P(\varepsilon) - P(0)$ as a negligible (i.e. regular) perturbation of F in the sense of (3). We plan to build on such insights in order to develop a constructive procedure for determining accurate time scale decompositions of systems as in (1.11).

(6) There are numerous ties to other research areas -- almost-invariant subspaces and implicit systems (i.e. systems of the form $E\dot{x} = Ax$), to name two -- which we feel can be illuminated significantly using the algebraic framework we have developed.

(7) We have also begun to use the results obtained to date to examine more closely the conditions $A(\varepsilon)$ must satisfy in order for there to exist an approximation as in (1.2). In particular, there are two ways in which a system can fail to satisfy the conditions. Either $A(\varepsilon)$ violates the so-called multiple semisimple nullstructure (MSSNS) condition [2] or it satisfies this but violates the multiple semistability condition (MSST). If the first of these is violated it means that in the procedure for constructing the $A_i$ in (1.2) (either as in [1] or via Schur complementation as in [7], [9]) one encounters one of these matrices

which has a Jordan block of size greater than $1 \times 1$ corresponding to the zero eigenvalue. For example,

$$A(\varepsilon) = \begin{bmatrix} -\varepsilon & 1 \\ -\varepsilon & -\varepsilon \end{bmatrix} \tag{1.13}$$

is asymptotically stable for any $\varepsilon > 0$ but it does <u>not</u> have a uniform approximation as in (1.2) Note that the Jordan form of $A(\varepsilon)$ changes abruptly at $\varepsilon = 0$ which indicates that there must exist a singularity at $\varepsilon = 0$ in the similarity transformation that brings $A(\varepsilon)$ to Jordan form. For this reason we have begun to study alternative canonical forms of $A(\varepsilon)$ and in particular the relationship between the Smith decomposition and the eigenstructure of $A(\varepsilon)$. A conjecture which we are in the process of examining is that the eigenvalues of $A(\varepsilon)$ are of orders of $\varepsilon$ exactly equal to the invariant factors if and only if $A(\varepsilon)$ has MSSNS. This fact, if true, will also be of great value in several control problems we are considering.

The point we have just made -- that a violation of the MSSNS condition corresponds to singularities in the similarity transformation that brings $A(\varepsilon)$ to Jordan form -- suggests another related problem on which we have made progress recently. Specifically this problem is concerned with using $\varepsilon$-dependent similarity transformations of $A(\varepsilon)$ which <u>are</u> singular at 0 and cancel the singularities in the Jordan similarity transformation. That is, if $A(\varepsilon)$ does not have MSSNS, we are concerned with constructing a transformation $T(\varepsilon)$ so that

$$\hat{A}(\varepsilon) = T(\varepsilon)A(\varepsilon)T^{-1}(\varepsilon) \tag{1.14}$$

<u>does</u> have MSSNS. What this in essence does is identify those components of $x(t)$ which require scaling, an idea which has been used by Sannuti

and Wason[†]. Again our work to date indicates that the algebraic approach we have developed provides precisely the correct framework for answering this question in the simplest and most illuminating fashion. As an example, consider

$$A(\varepsilon) = \begin{bmatrix} -\varepsilon & 1 \\ 0 & -\varepsilon \end{bmatrix} \qquad (1.15)$$

If we scale the state

$$\hat{x} = T(\varepsilon)x = \begin{bmatrix} \varepsilon & 0 \\ 0 & 1 \end{bmatrix} x \qquad (1.16)$$

we find that

$$\hat{A}(\varepsilon) = \begin{bmatrix} -\varepsilon & \varepsilon \\ 0 & -\varepsilon \end{bmatrix} \qquad (1.17)$$

which <u>does</u> satisfy the MSSNS condition. Note that the invariant factors of $A(\varepsilon)$ are 1 and $\varepsilon^2$, while both invariant factors of $\hat{A}(\varepsilon)$ are $\varepsilon$, indicating this this system has only one time scale.

In [11] we have considered an alternative approach to the problem of deriving approximations for a particular class of systems which violate the MSSNS condition. In particular in this paper we examine a specific model structure commonly used in analyzing interconnected power systems. Specifically we have considered models of the form

---

[†] P. Sannuti and H. Wason, "Singular Perturbation Analysis of Cheap Control Problems," Proc. 22nd IEEE Conf. on Decision and Control, San Antonio, Texas, Dec. 1983

P. Sannuti and H. Wason, Int. J. Contr., Vol. 37, 1983, pp. 1259-1286.

$$A(\varepsilon) = \begin{bmatrix} 0 & F(\varepsilon) \\ I & 0 \end{bmatrix}$$

where $F(\varepsilon)$ is an infinitesimal stochastic matrix. Because of this, $F(\varepsilon)$ has a fixed zero eigenvalue and thus $A(0)$ doesn't have semisimple null structure. However, using the results in [2] on aggregation of finite-state Markov processes we are able to derive an approximation which in essence corresponds to keeping the dominant term is each element of $\exp\{A(\varepsilon)t\}$. For the class of systems considered in [11] this can be accomplished in a relatively simple and intuitively appealing fashion. We are at present considering generalizations to other systems, and the development of a precise definition of the way in which one should think of this approximation as being good. For example, $\hat{x}(t) = e^{t}$ is a "good" approximation of $x(t) = e^{(1+\varepsilon)t}$ in the sense that the coefficient multiplying t in the exponent of

$$\frac{x(t)}{\hat{x}(t)}$$

is of strictly higher order in $\varepsilon$ (which is not true of $e^{(0.9)t}$ or $e^{(1.1)t}$, for example).

The second way in which a system can fail to have an approximation as in (1.2) is if it satisfies the MSSNS condition but not the MSST condition. This could happen for one of two reasons. One possibility is that there are unstable poles such as $(\varepsilon + \varepsilon^2)$. The leading term approximation described in the preceding paragraph is aimed at such a situation. The other possibility is that $A(\varepsilon)$ is stable for $\varepsilon > 0$ but some of the eigenvalues of one of the $A_i$ in (1.2) are purely imaginary. This corresponds to a situation in which the rate of oscillation in a complex mode is of lower order (and hence faster) than the damping. Consider for example

$$A(\varepsilon) = \begin{bmatrix} -\varepsilon & 1 \\ -1 & -\varepsilon \end{bmatrix} \tag{1.18}$$

which yields responses of the form $e^{-\varepsilon t}$ sin t.

What we are considering in such cases is the inclusion of higher-order terms in the asyptotic expansion, or, equivalently allowing the $A_i$ in (1.2) to violate the condition $A_i A_j = 0$. To see how such a decomposition might be obtained, consider $y(t) = P^{-1}(\varepsilon)x(t)$. Then

$$\dot{y}(t) = D(\varepsilon)\overline{A}(\varepsilon)y(t), \quad \overline{A}(\varepsilon) = Q(\varepsilon)P(\varepsilon) \tag{1.19}$$

Compare this to the process z(t) defined in (1.7). If we define the "correction process"

$$w(t) = e^{-D(\varepsilon)\overline{A}t}y(t) \tag{1.20}$$

we find that

$$\dot{w}(t) = [-D(\varepsilon)A + e^{-D(\varepsilon)At}D(\varepsilon)A(\varepsilon)e^{D(\varepsilon)At}]\ w(t) \tag{1.21}$$

and an investigation of the structure of the matrix in (1.2) should identify the desired higher-order corrections. As a very simple example, consider again (1.18). In this case

$$\dot{w}(t) = \begin{bmatrix} -\varepsilon & 0 \\ 0 & -\varepsilon \end{bmatrix} w(t) \tag{1.22}$$

Our present work along the lines just described is quite close to providing a general procedure for approximating dynamics of the form of (1.1) which violate the conditions for (1.2) to exist. Such a procedure would involve $\varepsilon$-dependent similarity transformations to obtain a transformed system which satisfies the MSSNS condition, leading-order approximations for unstable modes of systems which satisfy MSSNS but have complex poles with real parts of higher order than imaginary parts. As an example of a system which requires two of these steps, consider again $A(\varepsilon)$ in (1.13). If we scale $A(\varepsilon)$ as in (1.14) with

$$T(\varepsilon) = \begin{bmatrix} \varepsilon^{1/2} & 0 \\ 0 & 1 \end{bmatrix} \tag{1.23}$$

we obtain

$$\hat{A}(\varepsilon) = \begin{bmatrix} -\varepsilon & \varepsilon^{1/2} \\ -\varepsilon^{1/2} & \varepsilon \end{bmatrix} \tag{1.24}$$

which is essentially of the same form as in (1.18) (one need only identify $\varepsilon^{1/2}$ as the fundamental parameter rather than $\varepsilon$ and perform a simple time scaling).

## II. Control and Estimation

Our research in this area has had two major thrusts. The first of these builds directly on the tools decribed in the preceding section. Specifically we have focused attention on an examination of systems of the form

$$\dot{x}(t) = A(\varepsilon)x(t) + B(\varepsilon)u(t) \tag{2.1}$$

$$y(t) = C(\varepsilon)x(t) \tag{2.2}$$

Our ultimate aim is to develop a complete picture of how time-scales, weak couplings, and differences in the scales of controllability and observability of various components of the state and in the weightings of states and controls in the system design criterion interact in determining the structure of control designs. Our goal is to develop constructive procedures for designing hierarchical or decentralized control systems which take into account these scaling differences to achieve nearly optimal performance.

Notable contributions have been made on various aspects of this subject, but much remains to be done. Our results to date indicate that the algebraic approach outlined in the preceding section provides an excellent framework for examining this subject, for obtaining results which extend considerably what is known at present, and for shedding substantial light on the nature of problems of this type by uncovering and explicitly examining the critical mathematical constructs which form the heart of these problems. For example, as indicated in the preceding section, our results have shown that the invariant factors of $A(\varepsilon)$ determine the open-loop time scales of (2.1), assuming that the MSSNS condition is satisfied (if is is not, some scaling must be performed as described in Section I). Consequently a natural question to ask is to determine precisely how the invariant factors of (2.1) can be modified by feedback of the form

$$u(t) = K(\varepsilon)x(t) \tag{2.3}$$

(where $K(\varepsilon)$ is again a matrix over the ring of functions analytic at $\varepsilon=0$). By allowing $\varepsilon$-dependence in (2.3) we are in essence considering the question of feedback structure. That is, the matrix $K(0)$ determines which states are strongly coupled to which controls, the matrix

$$\frac{K(\varepsilon)-K(0)}{\varepsilon}$$

determines the next level of coupling in the hierarchy, etc.

At this time, we have obtained important results on invariant factor assignment [7], [9]. In particular if $A(\varepsilon)$ and $B(\varepsilon)$ are left coprime, i.e., if $[A(0) \vdots B(0)]$ has full row rank, then the closed-loop system matrix

$$F(\varepsilon) = A(\varepsilon) + B(\varepsilon)K(\varepsilon) \tag{2.4}$$

can have no more than $b$ = rank $B(0)$ non-unit invariant factors, and these $b$ factors can be made to equal an arbitrary set $\varepsilon^{j_1},\ldots,\varepsilon^{j_b}$ (with $\varepsilon^0 = 1$, $\varepsilon^\infty = 0$) by an appropriate choice of $K(\varepsilon)$ which we can construct explicitly. This result opens the way for the consideration of numerous other problems:

(1) Precisely how can the eigenvectors of $F(\varepsilon)$ be controlled as well as the invariant factors? That is, how can we influence which states evolve at which time scales?

(2) Can $K(\varepsilon)$ be chosen so that desired invariant factors are achieved <u>and</u> $F(\varepsilon)$ has MSSNS? If not, characterize the required scaling of $F(\varepsilon)$.

(3) Note that if $b$ = rank $B(0)$ < $m$ = rank $B(\varepsilon)$ for $\varepsilon>0$, our result indicates that fewer time scales can be affected than we have independent controls. In such a case, some of the controls are uniformly weak, and the only way in which time scales could be influenced in general is by high

gain -- i.e. by allowing terms of the form $1/\varepsilon$ in $K(\varepsilon)$ or equivalently by allowing input scaling

$$u(t) = S(\varepsilon)\hat{u}(t) \tag{2.5}$$

so that

$$\dot{x}(t) = A(\varepsilon)x(t) + \hat{B}(\varepsilon)\hat{u}(t) \tag{2.6}$$

where $\hat{B}(\varepsilon) = B(\varepsilon)S^{-1}(\varepsilon)$ is still analytic at $\varepsilon=0$ and has the property that $\hat{B}(0)$ is of full rank. For example, the time scale of

$$\dot{x} = -x + \varepsilon u \tag{2.7}$$

can be charged by using feedback of the form

$$u = (\frac{1}{\varepsilon} + K(\varepsilon))x \tag{2.8}$$

(4) Another important problem is the case when $[A(\varepsilon) \vdots B(\varepsilon)]$ does not have full row rank. In this case, there are two avenues of investigation. The first of these involves the use of scaling of the inputs and possibly the states to achieve the coprime and $B(0)$ full rank conditions. In our other approach, we suppose that we are restricted to using $K(\varepsilon)$ which are analytic at $\varepsilon=0$ (and thus perform no input scaling). In this case $F(\varepsilon)$ is of the form

$$F(\varepsilon) = W(\varepsilon)\overline{F}(\varepsilon),\ \overline{F}(\varepsilon) = \overline{A}(\varepsilon) + \overline{B}(\varepsilon)K(\varepsilon) \tag{2.9}$$

where $W(\varepsilon)$ is a greatest common left divisor of $A(\varepsilon)$, $B(\varepsilon)$, and $\overline{A}(\varepsilon)$, $\overline{B}(\varepsilon)$ are coprime. If the invariant factors of $F(\varepsilon)$, $W(\varepsilon)$, and $\overline{F}(\varepsilon)$ are denoted by $f_i(\varepsilon)$, $w_i(\varepsilon)$, and $\overline{f}_i(\varepsilon)$ and are ordered such that the ith one divides the (i+1)-th one, we have that

$$w_i(\varepsilon) \mid f_i(\varepsilon) \text{ and } \overline{f}_i(\varepsilon) \mid f_i(\varepsilon) \tag{2.10}$$

The first condition shows that every invariant factor of $F(\varepsilon)$ must contain the corresponding invariant factor of $W(\varepsilon)$. The $\bar{f}_i(\varepsilon)$ are governed by our result in the coprime case, and thus some conclusions about the $f_i(\varepsilon)$ can be drawn from the second divisibility condition in (2.10). Note that this does not provide a complete solution, and open questions remain. In particular, in [9] we present a result on one set of conditions under which $f_i(\varepsilon) = w_i(\varepsilon)\bar{f}_i(\varepsilon)$. Work on more complete characterizations of $f_i(\varepsilon)$ in other cases is continuing.

(5) There are close ties between our work and several other research areas which we have begun to explore and develop. In particular, questions such as (1) are related to the much broader subject of the geometric structure of (2.1), which in turn has ties to the work on almost (A,B) invariant subspaces of Willems†. In our case we have additional structure, however, provided by the various scales defined by increasing orders in $\varepsilon$. Also, just as the work of Willems has close ties to the topic of high gain feedback, so does our work, and we plan to explore this avenue. In particular we have begun to examine the interpretation and extension of the approach of Sannuti and Wason (referenced earlier) to our framework. As a final point, we note that it is certainly possible to consider choices of input scaling so that $\hat{B}(\varepsilon)$ has singularities at $\varepsilon=0$. This appears at least cosmetically to be more closely tied to work such as that on cheap control and high-gain feedback. However, if $\hat{B}(\varepsilon)$ contains terms of the form $1/\varepsilon^n$, a simple time scaling (so that the fastest time scale is the "new"

---

† J.C. Willems, "Almost Invariant Subspaces: An Approach to High Gain Feedback Design -- Part I: Almost Controlled Invariant Subspaces," IEEE Trans. on Aut. Control, Vol. AC-26, 1982, pp. 235-252.

time variable t) removes these singularities. Consequently our investigation will allow us to consider high gain by a simple time scale identification.

Another area [9] in which we have begun work is in the examination of a generalization of the cheap control problem which allows the open-loop system to have several time scales and allows differences in the scales at which different states and controls are weighted. Specifically, consider the problem of choosing a control law for (2.1) to minimize

$$J = \int_0^\infty [x'(t)Q(\varepsilon)x(t) + u'(t)R(\varepsilon)u(t)]dt \tag{2.11}$$

The Hamiltonian matrix for this problem is

$$H(\varepsilon) = \begin{bmatrix} A(\varepsilon) & -B(\varepsilon)R^{-1}(\varepsilon)B'(\varepsilon) \\ -Q(\varepsilon) & -A'(\varepsilon) \end{bmatrix} \tag{2.12}$$

Define $\Theta_F(\varepsilon)$ and $\Theta_b(\varepsilon)$ as the positive definite solutions of the algebraic Riccati equations

$$\Theta_f(\varepsilon)A(\varepsilon) + A'(\varepsilon)\Theta_F(\varepsilon) - \Theta_F(\varepsilon)B(\varepsilon)R^{-1}(\varepsilon)B'(\varepsilon)\Theta_F(\varepsilon) + Q(\varepsilon) = 0 \tag{2.13a}$$

$$\Theta_b(\varepsilon)A(\varepsilon) + A'(\varepsilon)\Theta_b(\varepsilon) + \Theta_b(\varepsilon)B(\varepsilon)R^{-1}(\varepsilon)B'(\varepsilon)\Theta_b(\varepsilon) - Q(\varepsilon) = 0 \tag{2.13b}$$

Then one can construct a similarity transformation

$$T(\varepsilon) = \begin{bmatrix} \Theta_F(\varepsilon) & -I \\ \Theta_b(\varepsilon) & -I \end{bmatrix} \tag{2.14}$$

operating on $H(\varepsilon)$ to yield

$$\hat{H}(\varepsilon) = \begin{bmatrix} -A'(\varepsilon) + \Theta_f(\varepsilon)B(\varepsilon)R^{-1}(\varepsilon)B'(\varepsilon) & 0 \\ 0 & -A'(\varepsilon) - \Theta_b(\varepsilon)B(\varepsilon)R^{-1}(\varepsilon)B'(\varepsilon) \end{bmatrix}$$

(2.15)

which indicates the well-known result that the eigenvalues of $H(\varepsilon)$ are also those of the optimal closed-loop system. This suggests that if $H(\varepsilon)$ has MSSNS, invariant factor analysis of $H(\varepsilon)$ may yield the time-scale structure of the closed-loop system. However, the similarity transformation bringing $H(\varepsilon)$ to the form (2.15) is unimodular if and only if $\Theta_F(\varepsilon) + \Theta_b(\varepsilon)$ is unimodular, which will not be the case in nearly singular control problems. This obviously points to the need for scaling and to the roles of $\Theta_F(\varepsilon)$, $\Theta_b(\varepsilon)$, and $H(\varepsilon)$ in determining the requisite scaling and the resulting time scale structure. Sannuti and Wason have investigated this point in the special case in which $R(\varepsilon) = \varepsilon R$ is the only $\varepsilon$-dependence (see also the closely related and important work of Hautus and Silverman[†]). We are now involved in examination of the general problem we have posed using the algebraic framework we have developed, and the extension of this problem to include $\varepsilon$-dependent observations in order to achieve our objective of developing a complete picture of the interplays among scales on open-loop dynamics, control effectiveness, observability, and weightings on inputs and states.

We have also made progress in our research involving estimation of finite-state Markov processes (FSMP's) possessing several time scales. The basis for this research is the methodology developed in [2] which uses our results on decomposing systems of the form (1.1) and the basic properties of FSMP's to construct a hierarchy of simpler, aggregated models of FSMP's which contain rare transitions. Each model ignores transitions that occur at a time scale far greater than the one with which the model is concerned and aggregates the effects of transitions that occur at faster

[†] M.L.J. Hautus and L.M. Silverman, "System Structure and Singular Control," Linear Algebra & Applications, Vol. 50, pp. 369-402, 1983.

scales. The existence of such a hierarchy suggests the use of estimator structures which take advantage of such a decomposition of the underlying process, thereby offering the possibility of reducing extremely complex estimation problems to sets of far simpler ones.

Our research in this area has consisted of two distinct pieces. One the one hand we have made progress in performing detailed asymptotic analyses of very simple singularly perturbed FSMP estimation problems [12] and this work has produced both several important insights into what types of performance measures are important for such estimators and an analytical approach for calculating asymptotic approximations to such measures. The other portion of our work [8], [10] has dealt directly with a class of FSMP's of great complexity but which also possess important structural features. Our objective in this area has been to develop estimator structures that take direct advantage of this structure. By doing so, it has been our hope to uncover important principles and concepts that could then be used both for designing estimators for other classes of problems and for suggesting promising and important theoretical directions.

The research described in [10] has as its motivation the automated analysis of electrocardiograms (ECG's). Our reason for choosing this focus is not only that ECG analysis is an important and challenging problem but also that it is necessary to establish a context for an investigation of this type. The class of "large and complex FSMP's" is far too amorphous to yield interesting insights and analysis; what is needed is to define a structured class of FSMP's with clear estimation objectives. Thus an accurate statement is that ECG analysis has guided the choices of estimation structures and problems we are investigating, but that the class of problems we are considering is by no means restricted to ECG analysis and includes numerous other complex signal analysis problems as well as topics such as multitarget tracking and complex queueing networks.

To be more specific, the problems we have been analyzing are hybrid in nature -- that is, they involve both discrete- and continuous-valued processes, where one can think of sequences of discrete states as events which influence the observed continuous waveforms. In particular, the type of model that we are considering consists of an interconnection of discrete-state processes where the state of one process can influence the transition rates in the other processes (as we will point out shortly, this is precisely how one can interpret the results of [2]), and particular transitions in some of these processes initiate the generation of continuous waveforms. The actual observation is the superposition of the continuous waveforms that have been generated (just as the ECG is the superposition of the measured electrical activity of the various regions the heart.

In [8], [10] a methodology is developed for modeling cardiac activity and in particular its effect on the observed ECG using models of this type. These models have several very important aspects. Two of these are <u>timing</u> and <u>control</u>. The issue of control is related to the fact that the electrical state of one portion of the heart -- represented by one of the finite-state processes in the model -- can strongly influence the future behavior of other portions of the heart. The issue of timing is concerned with the fact that one can observe dramatic differences in the influence the state of one portion of the heart can have on another, depending upon the state the other portion is in (see [8] and [10] for numerous examples). A third extremely important aspect of these cardiac models is that the time scale at which interactions among the discrete models change and at which continuous waveforms are initiated is <u>far slower</u> than the transition-by-transition scale at which each process evolves. It is this feature that suggests a decomposition of the estimator (which processes the observed ECG in order to track cardiac activity) in which the estimator for each subprocess has a highly aggregated model of the remainder of the overall process that is accurate enough at the coarse time scale at which it is important. This leads to estimation structures consisting of interconnections of discrete-state estimators which take as inputs the observed ECG <u>and</u> estimates from other local estimators and which produce estimates of state trajectories.

In the recent past we have been developing and analyzing estimators for processes that possess the features we have just described. Our analysis has been driven by concerns that differ from those which are usually considered in examining estimator performance but which are quite natural for discrete processes of the type we have described and in particular for the ECG problem. In particular, in usual estimation problems one measures performance by comparing the actual state and the estimate <u>at a particular point in time</u>. In discrete event-oriented problems such as ECG analysis one is more interested in the timing of events (especially those which determine the control behavior of the heart). Thus one is concerned with errors in <u>time corresponding to particular values of the state or state transition</u>. That is, an estimate $\hat{x}$ may be considered to be quite good even if $x(t) - \hat{x}(t)$ is often quite large if in fact the state and estimate trajectories have only small time shifts between them. A second important performance measure is error recovery, a concept that is most easily stated in coding terms. Specifically, if we think of the observations (e.g. the ECG) as an encoding of discrete events, then we would like our decoder (estimator) to have the property that the occurrence of inevitable decoding errors should <u>not</u> lead to long strings of subsequent decoding errors.

The other portion of our research on FSMP's [12] deals with the detailed asymptotic analysis of a simple FSMP estimation problem. While this problem is not rich enough to capture all aspects of the concerns described in the preceding paragraph, it has proved to be extremely useful in allowing us to begin to develop quantitative, analytic methods for problems of this type. Just as in the control work described earlier in this section, our fundamental interest in this problem is to understand how process time scale, observability (i.e. measurement information rates) and estimation criterion interact.

A first problem considered in [12] is the simple two-state process x(t) depicted in Figure 2.1 where $\lambda_1$ and $\lambda_2$ are of the same order of magnitude and where we have observations

$$dy(t) = h(x(t))dt + bdw(t) \tag{2.16}$$

where $E[dw^2(t)] = dt$. Letting $\pi_1 = \text{Prob}\ \{x(t) = 1|y(s),\ 0\leq s\leq t\}$ (2.17)

$$\pi_1 = \text{Prob}\ \{x(t) = 1|y(s),\ 0\leq s\leq t\}$$

we can write

$$\begin{aligned} d\pi_1(t) &= [-\lambda_1\pi_1(t) + \lambda_2(1-\pi_1(t))]dt \\ &\quad + \frac{1}{b^2}[h(t) - \bar{h}(\pi_1(t))][dy(t) - \bar{h}(\pi_1(t))dt] \end{aligned} \tag{2.18}$$

where

$$\bar{h}(\pi_1(t)) = h(1)\pi_1(t) + h(2)(1-\pi_1(t)) \tag{2.19}$$

As discussed in [12] there are four natural quantitative measures for the performance of this filter:

(1) Filter bias. This is the distance between the equilibrium value of $\pi_1(t)$ given that x(t) = 1 or 2 respectively and

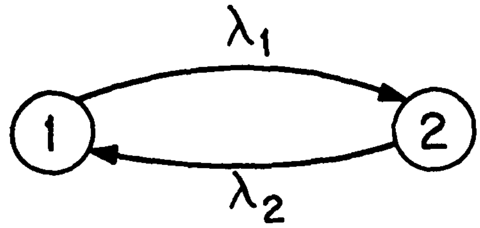



Figure 2.1

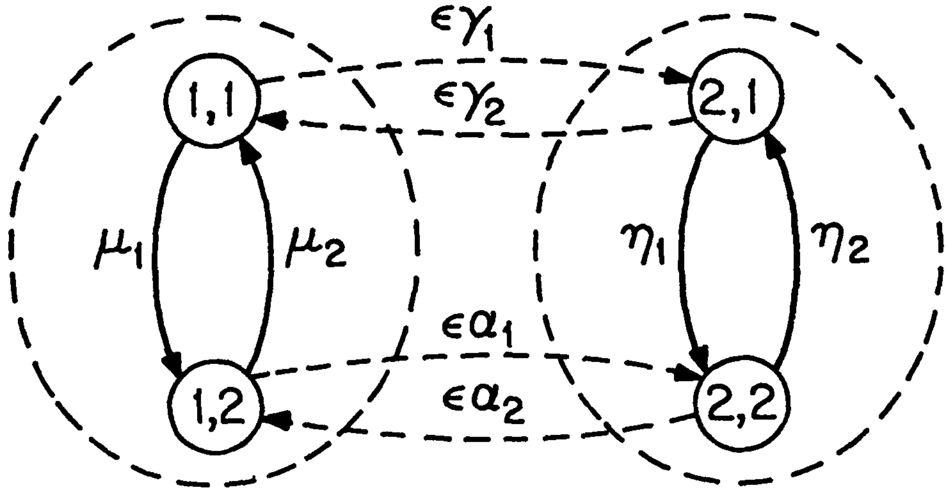



Figure 2.2

the corresponding boundary -- i.e. $\pi_1 = 1$ if $x(t) = 1$ or $\pi_1 = 0$ if $x(t) = 0$. This yields a measure of the ability to distinguish between the two states.

(2) Variance. The variance of deviations of $\pi_1$ around its equilibrium points excluding large deviations (i.e. false alarms). This corresponds most closely to the usual notion of estimation performance.

(3) Detection delays. The time it takes the filter to evolve from one equilibrium point to a detection threshold near the other equilibrium point following a transition in $x(t)$.

(4) Mean time between false alarms. The expected time between crossings of the threshold corresponding to the incorrect value of $\alpha(t)$ given that $x(t)$ has *not* changed.

All of these involve examining (2.18) assuming $x(t) = 1$ or $x(t) = 2$.

If $x(t) = 1$ over the interval of interest $\pi_1(t)$ evolves according to

$$d\pi_1(t) = [-\lambda_1\pi_1(t) + \lambda_2(1-\pi_1(t))]dt + K^2\pi_1(t)(1-\pi_1(t))^2 dt + K\pi_1(t)(1-\pi_1(t))dw(t) \tag{2.20}$$

where

$$K^2 \triangleq \left(\frac{h(t)-h(t)}{b}\right)^2 \tag{2.21}$$

is the rate at which information is accumulated that distinguishes between the hypotheses $x(t) = 1$ and $x(t) = 2$. If $x(t) = 2$ over the time interval, then

$$d\pi_1(t) = [-\lambda_1\pi_1(t) + \lambda_2(1-\pi_1(t))]dt - K^2\pi_1(t)^2(1-\pi_1(t))dt + K\pi_1(t)(1-\pi_1(t))dw(t) \tag{2.22}$$

Note that in either case $\pi_1(t)$ is a diffusion process on the bounded domain [0,1] and in fact the boundaries of this domain are so-called entrance boundaries.

The evaluation of the performance measures described previously involves the detailed study of (2.20) and (2.22). For example, the equilibria required to determine filter bias are the stationary points of deterministic systems obtained by setting the noise terms to zero in these equations. Note also that it is here that we begin to see a need for asymptotic analysis -- we have two rates, one determined by $K^2$ and one by $\lambda = (\lambda_1 + \lambda_2)/2$. It should therefore not be surprising that

$$\gamma = \frac{K^2}{\lambda} \tag{2.23}$$

(which roughly has the interpretation as the expected amount of information collected between x(t) Transitions ) is the critical quantity in evaluating asymptotic approximations to the performance measures just described. In particular, we have analyzed in detail the case where $\lambda = O(\varepsilon)$, i.e. where $\gamma$ is large $(O(\frac{1}{\varepsilon}))$. If we think of defining detection thresholds at values $\pi_1 = \delta$ and $\pi_1 = 1-\delta$, we have determined that $\delta$ must be chosen very carefully as a function of $\varepsilon$ in order to obtain detection delays that are small compared to the time between transitions and also to avoid catastrophic streams of false alarms. In particular a choice of $\delta(\varepsilon) = O(\sqrt{\varepsilon})$ leads to detection delays which go to infinity but at a much slower rate $((-\log\delta(\varepsilon))/K^2)$ than the time between transitions $(O(\frac{1}{\varepsilon}))$. Thus the estimator is correct "most of the time". Also, with this choice of threshold there will be O(1) false alarm between x(t) transitions.

The problem just described serves as a first step in analyzing the two-time-scale process $(x_1(t), x_2(t))$ illustrated in Figure 2.2 with measurements

$$dy_1(t) = h_1(x_1(t))dt + b_1 dw_1(t) \tag{2.24}$$

$$dy_2(t) = h_2(x_2(t))dt + b_2 dw_2(t) \tag{2.25}$$

First note that using the analysis in [2] this process can be decomposed into a hierarchy of two two-state process: a slow process corresponding to transitions in $x_1(t)$ and a fast process corresponding to $x_2(t)$. Note that this structure is exactly of the form under investigation in [8], [10] -- an interconnection of two processes in which rare transitions in the $x_1(t)$ process influence the transition rates ($\mu_1$ and $\mu_2$ vs. $\eta_1$ and $\eta_2$) of the fast process.

The structure of the estimator under investigation is the following:

(1) The measurements $y_1(t)$ are processed using the aggregate two-state Markov model for $x_1(t)$ evolving at the slow time scale. Since the difference between the actual evolution of $x_1(t)$ and that predicted by the approximate model is $O(\varepsilon)$, the conclusions described previously for the two-state process hold here as well.

(2) Given the estimate $\hat{x}_1(t)$, an estimate of $\hat{x}_2(t)$ is generated by using $y_2$ and the two-state model for $x_2(t)$ corresponding to $\hat{x}_1(t)$. Performance here can be evaluated in a fashion similar to that described for the process in Figure 2.1. Once can evaluate the performance when $\hat{x}_1$ is correct or in error, but the difference is significant only if the $\mu$'s and $\eta$'s are of different orders of magnitude.

(3) The estimator structure described by (1), (2) is not nearly optimal, but performs well under a wide range of conditions. The reason for this suboptimality is that there may exist nonnegligible information in $y_2$ concerning $x_1$ -- whether this difference is significant or not depends on the size of the differences between the $\mu$'s and $\eta$'s. Note that even if this difference is of no major consequence for estimating $x_2$, it may be significant for estimating $x_1$, since $x_1$ changes at a far slower time scale (and thus information can be accumulated over a much longer time period). We are presently completing our analysis of how the information in $y_2$ can be incorporated into the estimation of $x_1$. The basic idea is the following. Let $\hat{h}_2(x_1)$ be defined as the expected value of $h_2(x_2(t))$ given that $x_1$ is the correct value and $x_2(t)$ has reached its ergodic distribution. That is

$$\hat{h}_2(t) = \frac{\mu_2}{\mu_1+\mu_2} h_2(1) + \frac{\mu_1}{\mu_1+\mu_2} h_2(2) \qquad (2.26)$$

$$\hat{h}_2(2) = \frac{\eta_2}{\eta_1+\eta_2} h_2(1) + \frac{\eta_1}{\eta_1+\eta_2} h_2(2) \qquad (2.27)$$

Then the measurement $y_2$ can be written as

$$dy_2(t) = h_2(x_1(t))dt + [h_2(x_2(t)) - h_2(x_1(t))]dt + b_2dw_2(t) \quad (2.28)$$

Intuitively, if the information rate

$$\left[\frac{\hat{h}_2(1) - \hat{h}_2(2)}{b_2}\right]$$

is comparable to or greater than

$$\left[\frac{h_1(1) - h_1(2)}{b_2}\right]^2$$

one would expect $y_2$ to be of value in estimating $x_1$. Furthermore, if the conditional distribution of $x_1$ evolves at a slower time scale than the process $x_2$, one would expect that the term in brackets in (2.28) is negligible as far as $x_1$ estimation is concerned (although it is <u>all-important</u> as far as the estimation of $x_2$ is concerned!). In this case, we can use the approximation

$$dy_2(t) = \hat{h}_2(x_1(t))dt + b_2dw_2(t) \quad (2.29)$$

for the $x_1$-estimator which then has a form analogous to (2.18) except that it is driven by <u>both</u> observations.

## PUBLICATIONS

The publications listed below represent papers and reports supported in whole or in part by the Air Force Office of Scientific Research under Grant AFOSR-82-0258:

1. M. Coderch, A.S. Willsky, S.S. Sastry, and D.A. Castanon, "Hierarchical Aggregation of Linear Systems with Multiple Time Scales," IEEE Trans. on Automatic Control, Vol. AC-28, No. 11, Nov. 1983, pp. 1017-1030.

2. M. Coderch, A.S. Willsky, S.S. Sastry, and D.A. Castanon, "Hierarchical Aggregation of Singularly Perturbed Finite State Markov Processes," Stochastics, Vol. 8, 1983, pp. 259-289.

3. D.A. Castanon, M. Coderch, A.S. Willsky, "Hierarchical Aggregation of Diffusion Processes with Multiple Equilibrium Points," Proc. of the 1982 IEEE Conf. on Dec. and Control; extended version in preparation.

4. M. Coderch, "Multiple Time Scale Approach to Hierarchical Aggregation of Linear Systems and Finite State Markov Processes," M.I.T. Lab. for Inf. and Dec. Sys. Rept. LIDS-TH-1221, August 1982.

5. S.S. Sastry, "The Effects of Small Noise on Implicity Defined Non-linear Dynamical Systems," M.I.T. Lab. for Inf. and Dec. Sys. Rept. LIDS-P-1249; submitted for publication.

6. J.J. Slotine and S.S. Sastry, "Tracking Control of Non-linear Systems Using Sliding Surfaces with Application to Robot Manipulators," M.I.T. Lab. for Inf. and Dec. Sys. Rept. LIDS-P-1264; to appear in the International Journal on Control

7. X.-C. Lou, G.C. Verghese, A.S. Willsky, and M. Vidyasagar, "An Algebraic Approach to Analysis and Control of Time Scales," accepted for presentation at the 1984 American Control Conference, San Diego, Calif., June 1984.

8. P.C. Doerschuk, R.R. Tenney, and A.S. Willsky, "Estimation-Based Approaches to Rhythm Analysis in Electrocardiograms," Proc. of C.N.R.S. Meeting on Abrupt Changes in Signals and Systems," Paris, France, March 1984.

9. X.-C. Lou, "An Algebraic Approach to Analysis and Control of Time Scales," Ph.D. thesis proposal, to be submitted Feb. 1984.

10. P.C. Doerschuk, "Large-Scale System Estimation Problems and Electrocardiogram Interpretation," Ph.D. thesis proposal, Nov. 1983.

11. M. Coderch, A.S. Willsky, and S.S. Sastry, "Time Scale Analysis and Coherence Areas in Power Systems," in preparation.

12. M. Coderch and A.S. Willsky, "Asymptotic Analysis of Estimator Behavior for a Singularly-Perturbed Finite-State Markov Process," in preparation.

END

FILMED

6-84

DTIC